Available online at www.sciencedirect.com

ScienceDirect

International Journal of Approximate Reasoning
44 (2007) 339–357

INTERNATIONAL JOURNAL OF APPROXIMATE REASONING

www.elsevier.com/locate/ijar

Extreme points of coherent probabilities in finite spaces

Anton Wallner

Department of Statistics, Ludwig-Maximilians-University Munich, Ludwigstr. 33, D-80539 München, Germany

Received 15 December 2005; received in revised form 30 June 2006; accepted 31 July 2006
Available online 25 September 2006

Abstract

Every coherent probability (= F-probability) $\mathscr{F}$ on a finite sample space Ω_k with k elements defines a set of classical probabilities in accordance with the interval limits. This set, called "structure" of $\mathscr{F}$, is a convex polytope having dimension $\leqslant k-1$. We prove that the maximal number of extreme points of structures is exactly $k!$.

Keywords: Interval probability; Coherent probability; F-probability; Structure; Core; Credal set; Polyhedron; Polytope; Extreme point; Vertex; 0/1-Matrix

1. Introduction

There is a famous subclass of *coherent probabilities* = *F-probabilities* (terminologies of Walley [7] and Weichselbeirger [9] respectively) on finite sample spaces, for which the computing and counting of the *extreme points* = *vertices* of the corresponding *structures* are algorithmically very easy: the class of *2-monotone capacities* (= *C-probabilities* in Weichselberger's terminology). It is well-known that the structures of 2-monotone capacities on the sample space Ω_k with k elements have at most $k!$ vertices and that this bound is sharp.

Inspired by this and by the visible barycentric representation of structures for $k = 1, 2, 3, 4$ (where the classes of C- and F-probabilities coincide for $k = 1, 2, 3$) Weichselberger

E-mail address: toni@stat.uni-muenchen.de

0888-613X/$ - see front matter
doi:10.1016/j.ijar.2006.07.017

expressed the following conjecture in the early 1990s, i.e., at the beginning of his theory of *interval probability*:

Weichselberger conjecture (WEC). *The structures of all F-probabilities on Ω_k have at most $k!$ vertices.*

For more than one decade many efforts were made to find a counterexample or to verify Weichselberger's conjecture.

We here prove him right.

For everyone who is a little bit familiar with coherent (= F-) probabilities and the description of vertices (= extreme points) of polyhedra in general, Sections 2–4 are well-known as regards content, but not as regards the special nomenclature used here for the sake of convenience.

Following Weichselberger's terminology, Section 2 formally defines *F-probability*, *C-probability* and *structure*. Section 3 is concerned with basic facts about polyhedra and their vertices. Section 4 embeds structures of F-probabilities on Ω_k isomorphically into the space $\mathbb{R}^k$, for convincing that they "are in fact polytopes" which are, moreover, defined by 0/1-matrices, and for preparing that apart from this the F-property of F-probabilities is not needed to obtain the result.

However, the essential part is Section 5, where the (WEC) is proven.

Section 6 is reserved for concluding remarks.

2. F-probabilities, C-probabilities, and their structures

Here we report very briefly the main concepts of Weichselberger's theory of interval probability (see, e.g., [9]), in particular the concepts of *F-* and *C-probability* and, indispensably, of their *structures*.

Just as in classical probability theory, the starting point is some fixed *measurable space* $(\Omega;\mathscr{A})$, where Ω is a non-empty *sample space* and $\mathscr{A}$ is a *σ-algebra over* Ω. Here we only want to deal with the finite case. Hence we assume that Ω is finite, w.l.o.g. $\Omega = \Omega_k := \{1,\dots,k\}$ with $k \geqslant 1$, and $\mathscr{A} = \mathscr{P}(\Omega_k)$, i.e., the *power set of* Ω_k. The members A of $\mathscr{P}(\Omega_k)$ are called *events*, and, in particular, the singletons $E_i := \{i\}$, $i = 1,\dots,k$, are called *atoms*.

Definition 2.1.

(a) A set function $p\colon \mathscr{P}(\Omega_k) \to [0;1]$ is called a *K-function* (or *classical probability*) *on* $(\Omega_k;\mathscr{P}(\Omega_k))$, if it satisfies *Kolmogorov's axioms*, i.e.,
   (i) $p(A) \geqslant 0$, $\forall A \subseteq \Omega_k$;
   (ii) $p(\Omega_k) = 1$ (*norm condition*);
   (iii) p is *additive*, i.e., $p(A \cup B) = p(A) + p(B)$, $\forall A, B \subseteq \Omega_k$ with $A \cap B = \emptyset$.
   The set of all K-functions on $(\Omega_k;\mathscr{P}(\Omega_k))$ is denoted by $\mathscr{K}_k$.

(b) A quadruple $\mathscr{F} = (\Omega_k;\mathscr{P}(\Omega_k);L,U)$ is called an *F-(probability) field on* $(\Omega_k;\mathscr{P}(\Omega_k))$, if L, U: $\mathscr{P}(\Omega_k) \to [0;1]$ are set functions such that

$$\mathscr{M}(\mathscr{F}) := \{p \in \mathscr{K}_k \mid L(A) \leqslant p(A) \leqslant U(A), \forall A \subseteq \Omega_k\}$$

is not empty and $L(A) = \min_{p\in\mathscr{M}(\mathscr{F})} p(A)$, $U(A) = \max_{p\in\mathscr{M}(\mathscr{F})} p(A)$, $\forall A \subseteq \Omega_k$.

(c) The set $\mathscr{M}(\mathscr{F})$ according to (b) is named the *structure of* $\mathscr{F}$.

Sometimes structures are also called *cores* (see, e.g., [5]) or *credal sets* (see, e.g., [3]).

There are four basic properties of F-fields:

Corollary 2.2. *Let $\mathscr{F} = (\Omega_k; \mathscr{P}(\Omega_k); L, U)$ be an F-field. Then we have:*

(a) $L(.) \geqslant 0$.
(b) $L(\emptyset) = 0$.
(c) $L(\Omega_k) = 1$ (*i.e.*, L *is* normed).
(d) $U(A) = 1 - L(\Omega_k \backslash A)$, $\forall A \subseteq \Omega_k$.

Proof. (a), (b) and (c): straightforward. For (d) use $p(A) = 1 - p(\Omega_k \backslash A)$, $\forall A \subseteq \Omega_k$, $\forall p \in \mathscr{K}_k$. □

It is possible to describe the structure of an F-field only by its lower limit L:

Corollary 2.3. *Let $\mathscr{F} = (\Omega_k; \mathscr{P}(\Omega_k); L, U)$ be an F-field. Then*

$$\mathscr{M}(\mathscr{F}) = \{p \in \mathscr{K}_k \,|\, p(A) \geqslant L(A), \forall A \subseteq \Omega_k\} \tag{1}$$

already determines the structure of $\mathscr{F}$.

Proof. Straightforwardly with Corollary 2.2(d). □

As a special subclass of F-probabilities we make a note of C-probabilities = 2-monotone capacities:

Definition 2.4. A quadruple $\mathscr{C} = (\Omega_k; \mathscr{P}(\Omega_k); L, U)$ is called a *C-(probability) field on* $(\Omega_k; \mathscr{P}(\Omega_k))$, if the following conditions are satisfied:

(a) $L(.) \geqslant 0$.
(b) $L(\emptyset) = 0$.
(c) $L(\Omega_k) = 1$.
(d) $U(A) = 1 - L(\Omega_k \backslash A)$, $\forall A \subseteq \Omega_k$.
(e) L is 2-*monotone*, i.e., $L(A) + L(B) \leqslant L(A \cap B) + L(A \cup B)$, $\forall A, B \subseteq \Omega_k$.

Lemma 2.5. *Every C-field on $(\Omega_k; \mathscr{P}(\Omega_k))$ is an F-field on $(\Omega_k; \mathscr{P}(\Omega_k))$.*

Proof. For example, see [2]. □

The consideration of F- and C-fields, in particular the geometrical properties of their structures, will be continued in Section 4 – after fixing some basic facts about polyhedra and their vertices.

3. Basic facts about polyhedra and their vertices; the psi-function

In this section we give the usual definitions of polyhedra, polytopes, and their vertices and, moreover, introduce some notational conventions.

Conventions 3.1.

(a) For every finite set X let $|X|$ be the cardinality of X.
(b) The components of $x \in \mathbb{R}^k$ are denoted by $x(1), \ldots, x(k)$.
(c) For $x, y \in \mathbb{R}^k$ let $\langle x, y\rangle$ be the usual scalar product of x and y, i.e., $\langle x, y\rangle := \sum_{i=1}^{k} x(i)y(i)$.

Definition 3.2. Let $X \subseteq \mathbb{R}^k$.

(a) X is called a *polyhedron*, if X is a finite intersection of closed affine halfspaces of $\mathbb{R}^k$, i.e., if there exist $n \in \mathbb{N}$, $a_1, \ldots, a_n \in \mathbb{R}^k$, and $\beta_1, \ldots, \beta_n \in \mathbb{R}$ such that

$$X = \bigcap_{i=1}^{n} \{x \in \mathbb{R}^k | \langle a_i, x\rangle \geqslant \beta_i\}. \tag{2}$$

(b) A polyhedron X is called a *polytope*, if X is bounded.
(c) A point x of a polyhedron X is called a *vertex* or *extreme point of* X, if $\forall y, z \in X$ $\forall \lambda \in]0;1[$ $(x = (1-\lambda)y + \lambda z \Rightarrow x = y = z)$. The set of all vertices of X is denoted by $\mathscr{E}(X)$.

There is a well-known – but generally very ineffective – procedure for finding all the vertices of some given polyhedron. We describe it in the following lemma, point (a).

Lemma 3.3. *Let $X \subseteq \mathbb{R}^k$ be a polyhedron of the form (2).*

(a) *If $x \in X$, then x is a vertex of X iff there are indices $i_1, \ldots, i_k \in \{1, \ldots, n\}$ such that $\{a_{i_1}, \ldots, a_{i_k}\}$ is a basis of $\mathbb{R}^k$ and $\langle a_{i_j}, x\rangle = \beta_{i_j}$, $\forall j = 1, \ldots, k$.*
(b) $|\mathscr{E}(X)| \leqslant \binom{n}{k}$.

Proof. For (a) see, e.g., [4], 7.2(b), p. 122; (b) follows from (a). □

The estimation in Lemma 3.3(b) has the big disadvantage that it is very weak in general, but it has the advantage that it does not depend on $a_1, \ldots, a_n$ and $\beta_1, \ldots, \beta_n$. Subsequently we steer a middle course: We fix $a_1, \ldots, a_n \in \mathbb{R}^k$ and ask for the *maximal number of vertices* the corresponding polyhedra can have, if we let $\beta_1, \ldots, \beta_n$ vary over all (!) real numbers, i.e., if we allow all parallel translations of the defining hyperplanes.

For approaching this problem in a terminologically convenient manner, we now take a closer look at the definition (2) of a polyhedron X, which will lead to specific notation.

First, it is usual to employ the matrix-vector-style for an equivalent description of the polyhedron X: X is the solution space of the system

$$A \cdot x \geqslant b, \tag{3}$$

where $A := \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$ and $b := \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix}$, if we consider $a_1, \ldots, a_n$ as row vectors and x as a column vector in $\mathbb{R}^k$. But subsequently we do not need the description (3). Instead it will be more comfortable to use a functional style which is, moreover, close to the "language" of interval probability: Since w.l.o.g. we can assume that $a_1, \ldots, a_n$ are different and since the order of $a_1, \ldots, a_n$ is irrelevant, the relation $a_i \mapsto \beta_i$, $i = 1, \ldots, n$, is a function, and it is possible to represent X equivalently by

$$\{x \in \mathbb{R}^k | \langle a, x \rangle \geqslant L(a), \forall a \in \mathscr{A}_{\geqslant}\}, \tag{4}$$

if we put $\mathscr{A}_{\geqslant} := \{a_1, \ldots, a_n\}$ and $L\colon \mathscr{A}_{\geqslant} \to \mathbb{R}$, $L(a_i) := \beta_i$, $i = 1, \ldots, n$ (L stands for *lower bound*).

Secondly, of course it is feasible to incorporate into (4) linear equations $\langle a, x \rangle = L(a)$ by letting a, $-a \in \mathscr{A}_{\geqslant}$ and writing $\langle a, x \rangle \geqslant L'(a) \wedge \langle -a, x \rangle \geqslant L'(-a)$, where $L'(a) := L(a)$ and $L'(-a) := -L(a)$ (w.l.o.g. $L(a) = 0$, if $a = 0$, because we can presume that X is not empty). But, in order to avoid boring transformations when dealing with linear equations, we replace (4) formally by

$$\{x \in \mathbb{R}^k | \langle a, x \rangle \geqslant L(a), \forall a \in \mathscr{A}_{\geqslant}\} \cap \{x \in \mathbb{R}^k | \langle a, x \rangle = L(a), \forall a \in \mathscr{A}_{=}\}$$

for an appropriate set $\mathscr{A}_{=} \subseteq \mathbb{R}^k$. Moreover, it is handy to assume $\mathscr{A}_{=} \subseteq \mathscr{A}_{\geqslant}$.

We summarize:

Definition 3.4. Let $\mathscr{A}_{\geqslant} \subseteq \mathbb{R}^k$ be finite, $L\colon \mathscr{A}_{\geqslant} \to \mathbb{R}$, and $\mathscr{A}_{=} \subseteq \mathscr{A}_{\geqslant}$. Then we define:

(a) $\mathscr{M}(L) := \{x \in \mathbb{R}^k | \langle a, x \rangle \geqslant L(a), \forall a \in \mathscr{A}_{\geqslant}\}$.
(b) $\mathscr{M}(L, \mathscr{A}_{=}) := \mathscr{M}(L) \cap \{x \in \mathbb{R}^k | \langle a, x \rangle = L(a), \forall a \in \mathscr{A}_{=}\}$.

Corollary 3.5. *Let $\mathscr{A}_{\geqslant} \subseteq \mathbb{R}^k$ be finite and $L\colon \mathscr{A}_{\geqslant} \to \mathbb{R}$. Then we have:*

(a) $\mathscr{M}(L) = \mathscr{M}(L, \emptyset)$.
(b) $\mathscr{B}_{=} \subseteq \mathscr{A}_{=} \subseteq \mathscr{A}_{\geqslant} \Rightarrow \mathscr{E}(\mathscr{M}(L, \mathscr{A}_{=})) \subseteq \mathscr{E}(\mathscr{M}(L, \mathscr{B}_{=}))$.

Proof. (a) is trivial, and (b) follows from the fact that $\mathscr{M}(L, \mathscr{A}_{=})$ is a *face* of $\mathscr{M}(L, \mathscr{B}_{=})$, if $\mathscr{B}_{=} \subseteq \mathscr{A}_{=}$ (see, e.g., [4], p. 123ff). □

Definition 3.6. We define

$$\mathrm{BAS}(\mathscr{A}_{\geqslant}) := \{\mathscr{B} \subseteq \mathscr{A}_{\geqslant} | \mathscr{B} \text{ is a basis of } \mathbb{R}^k\}$$

for any $\mathscr{A}_{\geqslant} \subseteq \mathbb{R}^k$.

Using this abbreviation and the notations introduced in Definition 3.4 we can rewrite Lemma 3.3(a):

Corollary 3.7. *Let $\mathscr{A}_{\geqslant} \subseteq \mathbb{R}^k$ be finite, $L\colon \mathscr{A}_{\geqslant} \to \mathbb{R}$, and $\mathscr{A}_{=} \subseteq \mathscr{A}_{\geqslant}$. Then we have:*

(a) $\mathscr{E}(\mathscr{M}(L, \mathscr{A}_{=})) = \bigcup \{\mathscr{M}(L, \mathscr{A}_{=} \cup \mathscr{B}) | \mathscr{B} \in \mathrm{BAS}(\mathscr{A}_{\geqslant})\}$.
(b) $\mathscr{E}(\mathscr{M}(L)) = \bigcup \{\mathscr{M}(L, \mathscr{B}) | \mathscr{B} \in \mathrm{BAS}(\mathscr{A}_{\geqslant})\}$.

Proof. (a) can be deduced straightforwardly from Lemma 3.3(a), and (b) follows from (a) and Corollary 3.5(a). □

Note that in Corollary 3.7 all the sets $\mathscr{M}(L, \mathscr{A}_= \cup \mathscr{B})$ and $\mathscr{M}(L, \mathscr{B})$, $\mathscr{B} \in \mathrm{BAS}(\mathscr{A}_\geqslant)$, are empty or singletons (cf. Proposition 5.2).

Now we formally define the maximal number of vertices the polyhedra $\mathscr{M}(L, \mathscr{A}_=)$ can have, letting L vary over all real-valued functions with domain $\mathscr{A}_\geqslant$:

Definition 3.8. Let $\mathscr{A}_\geqslant \subseteq \mathbb{R}^k$ be finite and $\mathscr{A}_= \subseteq \mathscr{A}_\geqslant$. Then we define:

(a) $\psi(\mathscr{A}_\geqslant, \mathscr{A}_=) := \max\{|\mathscr{E}(\mathscr{M}(L, \mathscr{A}_=))| \,|\, L: \mathscr{A}_\geqslant \to \mathbb{R}\}$.
(b) $\psi(\mathscr{A}_\geqslant) := \psi(\mathscr{A}_\geqslant, \emptyset) = \max\{|\mathscr{E}(\mathscr{M}(L))| \,|\, L: \mathscr{A}_\geqslant \to \mathbb{R}\}$.

We refer to $\psi(\mathscr{A}_\geqslant, \mathscr{A}_=)$ and $\psi(\mathscr{A}_\geqslant)$ as the *psi-function of* $(\mathscr{A}_\geqslant, \mathscr{A}_=)$ or $\mathscr{A}_\geqslant$ respectively.

Therefore, both versions of the psi-function (itself) only depend on the dimension k of $\mathbb{R}^k$.

Corollary 3.9. *Let $\mathscr{A}_\geqslant \subseteq \mathbb{R}^k$ be finite. Then we have*:

(a) $\mathscr{B}_= \subseteq \mathscr{A}_= \subseteq \mathscr{A}_\geqslant \Rightarrow \psi(\mathscr{A}_\geqslant, \mathscr{A}_=) \leqslant \psi(\mathscr{A}_\geqslant, \mathscr{B}_=)$.
(b) $\mathscr{A}_= \subseteq \mathscr{A}_\geqslant \Rightarrow \psi(\mathscr{A}_\geqslant, \mathscr{A}_=) \leqslant \psi(\mathscr{A}_\geqslant)$.

Proof. (a) is a consequence of Corollary 3.5(b), and (a) implies (b). □

4. Making the goal smooth

In this section we will see that structures of F-probabilities on $(\Omega_k; \mathscr{P}(\Omega_k))$ are "in fact" polytopes in the space $\mathbb{R}^k$. Moreover, we will establish a relationship between their maximal vertices and both versions of the psi-function introduced in Section 3, which makes the goal, i.e., (WEC), "smooth" for its proof.

First, we interpret the elements p of $\mathscr{K}_k$ as vectors in $\mathbb{R}^k$ with components $p(E_1) = p(\{1\}), \ldots, p(E_k) = p(\{k\})$, and for brevity we write $p(1), \ldots, p(k)$. All the values $p(A)$, $A \subseteq \Omega_k$, are determined by these components due to the axiom of additivity:

$$p(A) = \sum_{i=1}^{k} 1_A(i) p(i), \quad \forall A \subseteq \Omega_k,$$

where 1_A is the *indicator function of* A, i.e.,

$$1_A(i) := \begin{cases} 1, & \text{if } i \in A, \\ 0, & \text{else.} \end{cases}$$

Indicator functions 1_A are also vectors in $\mathbb{R}^k$, only having components 0 or 1. Such vectors are called *0/1-vectors*. To simplify notation, we subsequently write A instead of 1_A. In particular, the empty event $\emptyset$ corresponds to $0 \in \mathbb{R}^k$, and Ω_k corresponds to $(1, \ldots, 1) \in \mathbb{R}^k$. Analogously, $\mathscr{P}(\Omega_k)$ can be seen as the set of all 0/1-vectors in $\mathbb{R}^k$, and the atoms $E_1, \ldots, E_k$ form the standard basis of $\mathbb{R}^k$. Note that $E_1, \ldots, E_k$ are also K-functions (*Dirac measures*), and the $(k-1)$-dimensional simplex spanned by those is exactly the set $\mathscr{K}_k$:

$$\mathscr{K}_k = \{p \in \mathbb{R}^k \,|\, \langle \Omega_k, p \rangle = 1 \wedge \langle E_i, p \rangle \geqslant 0, \forall i = 1, \ldots, k\}. \tag{5}$$

Now let $\mathscr{F} = (\Omega_k; \mathscr{P}(\Omega_k); L, U)$ be an F-field. If we rewrite Eq. (1) of its structure by

$$\mathscr{M}(\mathscr{F}) = \{p \in \mathbb{R}^k | \langle A, p \rangle \geqslant L(A), \forall A \in \mathscr{P}(\Omega_k)\} \cap \mathscr{K}_k$$

and take into account that $L(.)$ is non-negative and that $L(\Omega_k) = 1$ according to Corollary 2.2, (a) and (c), we infer from (5)

$$\mathscr{M}(\mathscr{F}) = \{x \in \mathbb{R}^k | \langle A, x \rangle \geqslant L(A), \forall A \in \mathscr{P}(\Omega_k)\} \cap \{x \in \mathbb{R}^k | \langle \Omega_k, x \rangle = L(\Omega_k)\}.$$

Using the terminology introduced in Definition 3.4(b) we conclude:

$$\mathscr{M}(\mathscr{F}) = \mathscr{M}(L, \{\Omega_k\}). \tag{6}$$

In particular, $\mathscr{M}(\mathscr{F})$ is a polyhedron in the space $\mathbb{R}^k$, and, since it is bounded by the simplex $\mathscr{K}_k$, $\mathscr{M}(\mathscr{F})$ is a polytope with dimension $\leqslant k-1$. For $k \leqslant 4$ it is possible to make structures visible using barycentric coordinates: cf. Figs. 1–4.

If we would prefer a matrix-vector-style for the description of $\mathscr{M}(\mathscr{F})$, we would have, let's say for $k = 4$, the following system of inequalities:

$$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} x(1) \\ x(2) \\ x(3) \\ x(4) \end{pmatrix} \geqslant \begin{pmatrix} L((0,0,0,0)) \\ L((1,0,0,0)) \\ L((0,1,0,0)) \\ L((0,0,1,0)) \\ L((0,0,0,1)) \\ L((1,1,0,0)) \\ L((1,0,1,0)) \\ L((1,0,0,1)) \\ L((0,1,1,0)) \\ L((0,1,0,1)) \\ L((0,0,1,1)) \\ L((1,1,1,0)) \\ L((1,1,0,1)) \\ L((1,0,1,1)) \\ L((0,1,1,1)) \\ L((1,1,1,1)) \end{pmatrix},$$

supplemented by the *norm condition*

$$\begin{pmatrix} 1 & 1 & 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} x(1) \\ x(2) \\ x(3) \\ x(4) \end{pmatrix} = L((1,1,1,1)). \tag{7}$$

Now let

$$M_k^F$$

be *the maximal number of vertices of structures of F-fields on* $(\Omega_k; \mathscr{P}(\Omega_k))$.

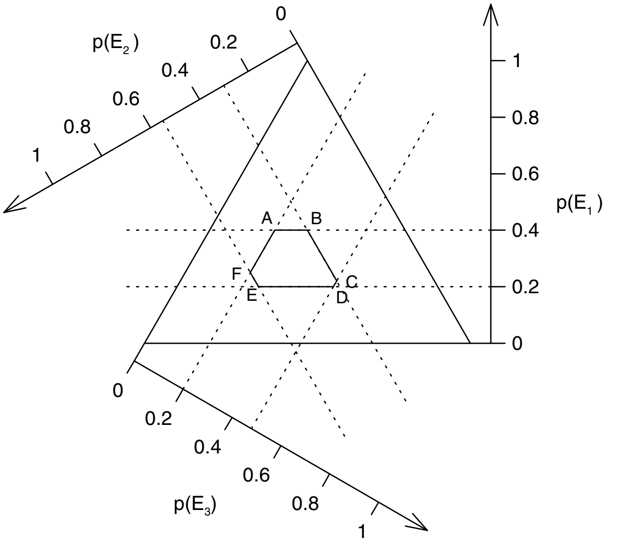

Fig. 1. An F-field $\mathscr{F} = (\Omega_3; \mathscr{P}(\Omega_3); L, U)$ using the barycentric representation. The probability intervals $P(.) = [L(.); U(.)]$ are given by

$$\begin{aligned} P(\emptyset) &= [0;0] & P(E_1 \cup E_2) &= [0.52; 0.80] \\ P(E_1) &= [0.20; 0.40] & P(E_1 \cup E_3) &= [0.45; 0.70] \\ P(E_2) &= [0.30; 0.55] & P(E_2 \cup E_3) &= [0.60; 0.80] \\ P(E_3) &= [0.20; 0.48] & P(\Omega_3) &= [1;1] \end{aligned}$$

Since $k = 3$, $\mathscr{F}$ is a C-field. Hence it is possible to describe the $3! = 6$ structure vertices A, B, C, D, E, and F by the characteristic Eqs. (9). For example, consider $p_D = (0.20, 0.32, 0.48) \in \mathscr{K}_3$ given by the point D: We have $p_D(E_1) = 0.20 = L(E_1)$, $p_D(E_1 \cup E_2) = 0.52 = L(E_1 \cup E_2)$, and $p_D(\Omega_3) = 1 = L(\Omega_3)$.

Our goal is to prove

$$M_k^F = k!.$$

First we say a word to the statement $M_k^F \geqslant k!$.

For this, let M_k^C be the maximal number of vertices of structures of C-fields on $(\Omega_k; \mathscr{P}(\Omega_k))$. By Lemma 2.5 we trivially get $M_k^F \geqslant M_k^C$. But for C-fields the maximal number of vertices, and moreover the description of the vertices itself, is well-known and easy (e.g., see [5], Theorem 3, p. 19, or [1], Proposition 13, p. 277):

Let $\mathscr{C} = (\Omega_k; \mathscr{P}(\Omega_k); L, U)$ be a C-field. Then, characteristically, we have $p \in \mathscr{E}(\mathscr{M}(\mathscr{C}))$ iff there exists a permutation π of $\{1, \ldots, k\}$ such that $p = p_\pi$, where p_π is defined by

$$p_\pi(E_{\pi(i)}) = L\left(\bigcup_{j=1}^{i} E_{\pi(j)}\right) - L\left(\bigcup_{j=1}^{i-1} E_{\pi(j)}\right) \tag{8}$$

for $i = 1, \ldots, k$, or – equivalently – by

$$p_\pi\left(\bigcup_{j=1}^{i} E_{\pi(j)}\right) = L\left(\bigcup_{j=1}^{i} E_{\pi(j)}\right) \tag{9}$$

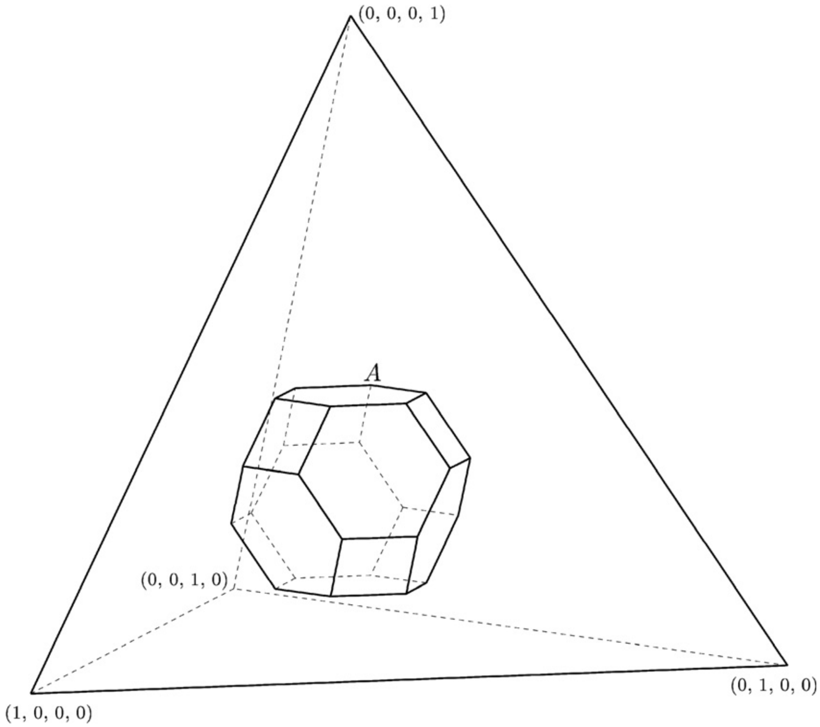

Fig. 2. An F-field $\mathscr{F} = (\Omega_4; \mathscr{P}(\Omega_4); L, U)$ using the barycentric representation. $\mathscr{F}$ is a C-field. For example, we verify (9) and (8) by considering the structure vertex p_A represented by the point A: p_A is determined by $p_A(E_1) = L(E_1)$, $p_A(E_1 \cup E_2) = L(E_1 \cup E_2)$, $p_A(E_1 \cup E_2 \cup E_3) = L(E_1 \cup E_2 \cup E_3)$, and $p_A(\Omega_4) = 1 = L(\Omega_4)$. Therefore

$$\begin{aligned} p_A &= (p_A(E_1), p_A(E_2), p_A(E_3), p_A(E_4)) \\ &= (L(E_1), L(E_1 \cup E_2) - L(E_1), L(E_1 \cup E_2 \cup E_3) - L(E_1 \cup E_2), 1 - L(E_1 \cup E_2 \cup E_3)). \end{aligned}$$

Here all the $4! = 24$ structure vertices of $\mathscr{F}$ can be obtained from p_A by permuting its coordinates. Hence $\mathscr{F}$ is a *uniform* C-field (cf. (10)).

for $i = 1, \ldots, k$ (cf. Figs. 1 and 2). (This correspondence between vertices and permutations is false for every F-field which is no C-field: cf. Figs. 3 and 4.) Hence, on the one hand, we get $M_k^C \leqslant k!$, but, on the other hand, there are C-fields which have exactly $k!$ structure vertices.

For example, consider a strict convex function f: $[0;1] \to [0;1]$ with $f(0) = 0$ and $f(1) = 1$, e.g., take $f(x) := x^2$, define L, U: $\mathscr{P}(\Omega_k) \to [0;1]$ by $L(A) := f\left(\frac{|A|}{k}\right)$ and $U(A) := 1 - L(\Omega_k \backslash A)$, $\forall A \subseteq \Omega_k$, and let $\mathscr{C} := (\Omega_k; \mathscr{P}(\Omega_k); L, U)$. Then it is easy to see that L is 2-monotone, and thus $\mathscr{C}$ is a C-field which, in addition, is *uniform* (cf. Fig. 2), i.e.,

$$L(A) = L(B), \quad \forall A, B \in \mathscr{P}(\Omega_k) \text{ with } |A| = |B|. \tag{10}$$

It is a straightforward exercise to show that for $\mathscr{C}$ defined like this the p_π's according to (8) are pairwise different. Hence $|\mathscr{E}(\mathscr{M}(\mathscr{C}))| = k!$ and thus $M_k^C = k!$.

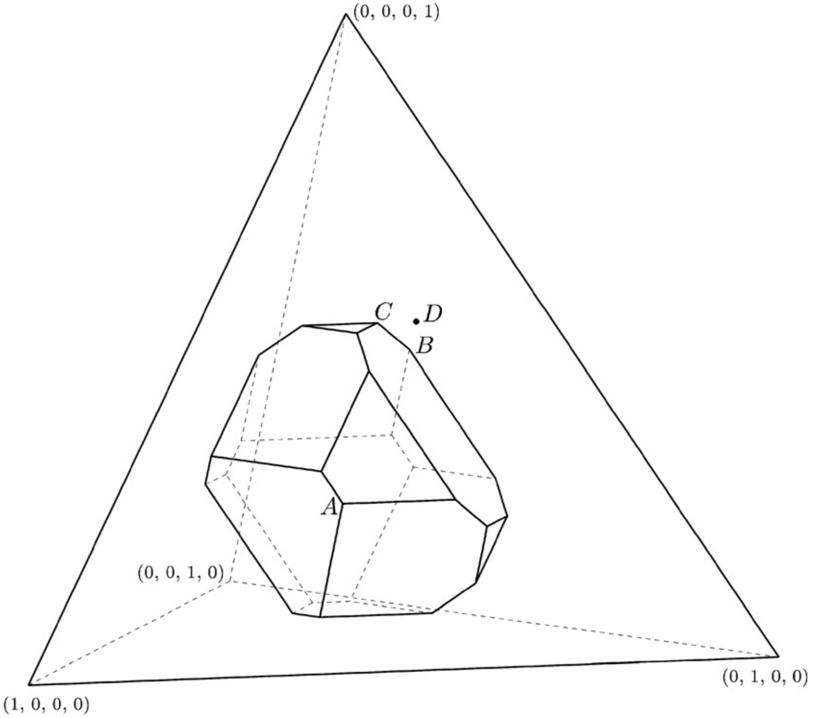

Fig. 3. An F-field $\mathscr{F} = (\Omega_4; \mathscr{P}(\Omega_4); L, U)$ using the barycentric representation. $\mathscr{F}$ has $4! = 24$ structure vertices, and as examples we consider the vertices p_A, p_B, and p_C corresponding to the points A, B, and C respectively:

- p_A is defined by $p_A(E_3) = L(E_3)$, $p_A(E_3 \cup E_4) = L(E_3 \cup E_4)$, $p_A(E_3 \cup E_4 \cup E_2) = L(E_3 \cup E_4 \cup E_2)$,
- p_B is defined by $p_B(E_1) = L(E_1)$, $p_B(E_1 \cup E_2) = L(E_1 \cup E_2)$, $p_B(E_1 \cup E_3) = L(E_1 \cup E_3)$,
- and p_C is defined by $p_C(E_1 \cup E_2) = L(E_1 \cup E_2)$, $p_C(E_1 \cup E_3) = L(E_1 \cup E_3)$, $p_C(E_1 \cup E_2 \cup E_3) = L(E_1 \cup E_2 \cup E_3)$,

supplemented by $p_A(\Omega_4) = p_B(\Omega_4) = p_C(\Omega_4) = 1 = L(\Omega_4)$. Since p_B and p_C do not respect the Eqs. (9), $\mathscr{F}$ is no C-field. Additionally, consider the point D: The corresponding K-function p_D, defined by $p_D\left(\bigcup_{j=1}^{i} E_j\right) = L\left(\bigcup_{j=1}^{i} E_j\right)$, $i = 1, 2, 3, 4$, lies outside the structure of $\mathscr{F}$.

We summarize:

Lemma 4.1. $M_k^F \geqslant k!$ *for all* $k \geqslant 1$.

Now we turn to the Weichselberger conjecture, i.e., to the statement

$$(\mathbf{WEC})\, M_k^F \leqslant k! \text{ for all } k \geqslant 1.$$

As mentioned in the Introduction, for $k = 1, 2, 3$ every F-field is a C-field; hence $M_k^F \leqslant k!$ is true at least for these three values of k. Moreover, in [8] it is shown (with completely different methods) that, for any $k \geqslant 1$, the bound $k!$ is valid for every *uniform* F-field (in the sense of (10), cf. Figs. 2 and 4).

Here we prove the entire (WEC). For preparing this, we propose two daring hypotheses. The *first one* is:

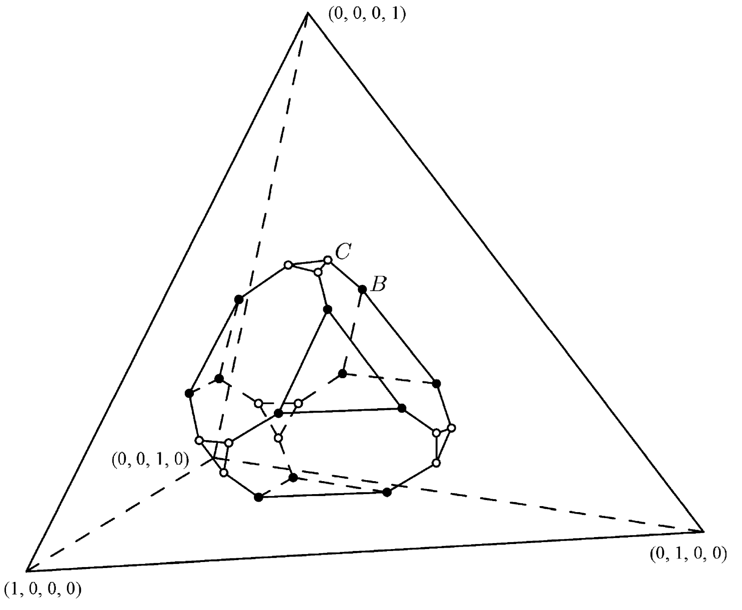

Fig. 4. A uniform (cf. (10)) F-field $\mathscr{F} = (\Omega_4; \mathscr{P}(\Omega_4); L, U)$ using the barycentric representation. The $4! = 24$ structure vertices of $\mathscr{F}$ can be divided into two parts: $\mathscr{E}_\bullet(\mathscr{M}(\mathscr{F}))$ and $\mathscr{E}_\circ(\mathscr{M}(\mathscr{F}))$ marked by $\bullet$ and $\circ$ resp. Let p_B, $p_C \in \mathscr{K}_4$ correspond to the points B and C resp. Then p_B and p_C are defined like p_B and p_C in Fig. 3. Hence $\mathscr{F}$ is no C-field. By the way, each of the convex hulls, of $\mathscr{E}_\bullet(\mathscr{M}(\mathscr{F}))$ and of $\mathscr{E}_\circ(\mathscr{M}(\mathscr{F}))$, are structures of uniform C-fields, each of them having "only a few" vertices (12), but together "spanning" the whole structure of $\mathscr{F}$ (see [8], Section 4.2, for more details on this subject).

- "The conditions on the lower bound L of an F-field are, compared with our problem, extremely weak. So we forget them at all."

This means: If we look at (6), we ignore the fact that the L in the term $\mathscr{M}(L, \{\Omega_k\})$ stems from an F-field. We consider the polytopes $\mathscr{M}(L, \{\Omega_k\})$ for *any* map L: $\mathscr{P}(\Omega_k) \to \mathbb{R}$, hoping that the maximal number of their vertices do not increase in comparison with the smaller class of L's which are lower bounds of F-fields.

In other words: Recalling the definition of the psi-function in 3.8(a) we make the following conjecture:

$$(\mathbf{WAC1})\ \psi(\mathscr{P}(\Omega_k), \{\Omega_k\}) \leqslant k! \text{ for all } k \geqslant 1.$$

Hence (WAC1) implies (WEC).

The *second hypothesis* is:

- "The *norm condition* (like (7) for $k = 4$) does not fit to the pattern of the remaining system of inequalities. So we forget it."

Geometrically this hypothesis says the following: Let, for fixed k, $L\colon \mathscr{P}(\Omega_k) \to \mathbb{R}$ be a function such that the corresponding polytope $\mathscr{M}(L, \{\Omega_k\})$ has the maximal number of vertices, i.e., $|\mathscr{E}(\mathscr{M}(L, \{\Omega_k\}))| = \psi(\mathscr{P}(\Omega_k), \{\Omega_k\})$. Then the (unbounded) polyhedron $\mathscr{M}(L)$

(without "norm condition") has no additional vertex, although $\mathscr{M}(L, \{\Omega_k\})$ is just a (lower dimensional) *face* of $\mathscr{M}(L)$ (cf. Corollary 3.5).

Again recalling the definition of the psi-function – now the second version in 3.8(b) –, the formal equivalent of the second hypothesis is the conjecture

(**WAC2**) $\psi(\mathscr{P}(\Omega_k)) \leqslant k!$ for all $k \geqslant 1$.

By Corollary 3.9(b) we have $\psi(\mathscr{P}(\Omega_k), \{\Omega_k\}) \leqslant \psi(\mathscr{P}(\Omega_k))$, and thus (WAC2) implies (WAC1).

In the next section we prove that indeed (WAC2) is true. But beforehand, we store for later reference:

Lemma 4.2. (WAC2) *implies* (WEC).

5. The main theorem

Now we concentrate on proving (WAC2). Let the natural number $k \geqslant 1$ be fixed. As already partially introduced in Section 4 we here officially (re-)define:

Definition 5.1.

(a) $a \in \mathbb{R}^k$ is called a *0/1-vector*, if $a(1), \ldots, a(k) \in \{0, 1\}$. The set of all 0/1-vectors in $\mathbb{R}^k$ is denoted by $\mathscr{P}(\Omega_k)$.
(b) Every element of $\mathrm{BAS}(\mathscr{P}(\Omega_k))$ (cf. Definition 3.6) is called a *0/1-basis*.

Clearly we have to deal with 0/1-vectors and 0/1-bases for proving (WAC2). But as far as possible and no additional work is needed, we will extend our view to all vectors in $\mathbb{R}^k$ and to all bases of $\mathbb{R}^k$.

We proceed in *four steps*, where our *first goal* is Lemma 5.3.

Proposition 5.2. *Let $\mathscr{B} = \{b_1, \ldots, b_k\}$ be a basis of $\mathbb{R}^k$ and $x, y \in \mathbb{R}^k$. Then we have*:

$$(\forall i = 1, \ldots, k.\, \langle b_i, x\rangle = \langle b_i, y\rangle) \quad \Rightarrow \quad x = y.$$

Proof. Well-known. □

Lemma 5.3. *Let $\mathscr{A}_\geqslant \subseteq \mathbb{R}^k$ be finite and $\mathscr{A} = \{a_1, \ldots, a_n\} \subseteq \mathscr{A}_\geqslant$. Let $\mathscr{B} = \{b_1, \ldots, b_k\} \subseteq \mathscr{A}_\geqslant$ be a basis of $\mathbb{R}^k$ such that there exist $\alpha_1, \ldots, \alpha_n \geqslant 0$ and $\beta_1, \ldots, \beta_k > 0$ with*

$$\sum_{i=1}^{n} \alpha_i a_i = \sum_{i=1}^{k} \beta_i b_i. \tag{11}$$

Then for all functions L: $\mathscr{A}_\geqslant \to \mathbb{R}$ and all $x, y \in \mathbb{R}^k$ we have (*cf.* Definition 3.4(*b*)):

$$x \in \mathscr{M}(L, \mathscr{A}) \wedge y \in \mathscr{M}(L, \mathscr{B}) \quad \Rightarrow \quad x = y.$$

Proof. Let all the premises be given. Then $x \in \mathscr{M}(L, \mathscr{A})$ implies

$$\langle a_i, x\rangle = L(a_i), \quad \forall i = 1, \ldots, n, \tag{12}$$

$$\langle b_i, x\rangle \geqslant L(b_i), \quad \forall i = 1, \ldots, k, \tag{13}$$

whereas $y \in \mathscr{M}(L, \mathscr{B})$ implies

$$\langle b_i, y\rangle = L(b_i), \quad \forall i = 1, \ldots, k, \tag{14}$$

$$\langle a_i, y\rangle \geqslant L(a_i), \quad \forall i = 1, \ldots, n. \tag{15}$$

Therefore we compute:

$$\begin{aligned}\sum_{i=1}^{n} \alpha_i L(a_i) &\overset{(12)}{=} \sum_{i=1}^{n} \alpha_i \langle a_i, x\rangle = \left\langle \sum_{i=1}^{n} \alpha_i a_i, x \right\rangle \overset{(11)}{=} \left\langle \sum_{i=1}^{k} \beta_i b_i, x \right\rangle = \sum_{i=1}^{k} \beta_i \langle b_i, x\rangle \\ &\overset{(13)}{\geqslant} \sum_{i=1}^{k} \beta_i L(b_i) \overset{(14)}{=} \sum_{i=1}^{k} \beta_i \langle b_i, y\rangle = \left\langle \sum_{i=1}^{k} \beta_i b_i, y \right\rangle \overset{(11)}{=} \left\langle \sum_{i=1}^{n} \alpha_i a_i, y \right\rangle \\ &= \sum_{i=1}^{n} \alpha_i \langle a_i, y\rangle \overset{(15)}{\geqslant} \sum_{i=1}^{n} \alpha_i L(a_i).\end{aligned}$$

Hence, at all positions we have in fact "=", in particular

$$\sum_{i=1}^{k} \beta_i \langle b_i, x\rangle = \sum_{i=1}^{k} \beta_i L(b_i).$$

From (13) and $\beta_1, \ldots, \beta_k > 0$ we conclude $\langle b_i, x\rangle = L(b_i)$, $\forall i = 1, \ldots, k$, hence, by (14), $\langle b_i, x\rangle = \langle b_i, y\rangle$, $\forall i = 1, \ldots, k$. Now Proposition 5.2 proves $x = y$. $\square$

The goal of the *second step* is Lemma 5.8, which can be understood as a condensation and interpretation of the result developed in Lemma 5.3: It is a step into the *dual space* of $\mathbb{R}^k$ connecting vertices in the primal space with simplices in the dual space.

Definition 5.4.

(a) For any set $X \subseteq \mathbb{R}^k$ let $\text{int}(X)$ be the *interior of* X (w.r.t. the usual topology on $\mathbb{R}^k$), i.e.,

$$\text{int}(X) := \{x \in \mathbb{R}^k \mid \exists \epsilon > 0.\, \mathrm{B}_\epsilon(x) \subseteq X\}, \quad \text{where} \quad \mathrm{B}_\epsilon(x) := \{y \in \mathbb{R}^k \mid \|x - y\| < \epsilon\}$$

and $\|.\|$ is some fixed norm on $\mathbb{R}^k$.

(b) For any finite set $\mathscr{A} = \{a_1, \ldots, a_n\} \subseteq \mathbb{R}^k$ let $\text{conv}(\mathscr{A})$ be the *convex hull of* $\mathscr{A}$, i.e.,

$$\text{conv}(\mathscr{A}) := \left\{ \sum_{i=1}^{n} \lambda_i a_i \,\middle|\, \lambda_1, \ldots, \lambda_n \geqslant 0, \sum_{i=1}^{n} \lambda_i = 1 \right\}.$$

(c) For every basis $\mathscr{B} = \{b_1, \ldots, b_k\}$ of $\mathbb{R}^k$ let

$\mathrm{S}(\mathscr{B}) := \text{conv}(\{0\} \cup \mathscr{B})$,

i.e., $\mathrm{S}(\mathscr{B})$ is the (k-dimensional) *simplex spanned by* $0, b_1, \ldots, b_k$.

Corollary 5.5. *For every basis $\mathscr{B} = \{b_1, \ldots, b_k\}$ of $\mathbb{R}^k$ we have*:

(a) $\mathrm{S}(\mathscr{B}) = \left\{\sum_{i=1}^k \lambda_i b_i \mid \lambda_1, \ldots, \lambda_k \geqslant 0,\ \sum_{i=1}^k \lambda_i \leqslant 1\right\}$.
(b) $\mathrm{int}(\mathrm{S}(\mathscr{B})) = \left\{\sum_{i=1}^k \lambda_i b_i \mid \lambda_1, \ldots, \lambda_k > 0,\ \sum_{i=1}^k \lambda_i < 1\right\}$.

Proof. (a) is trivial, and (b) can be deduced straightforwardly from (a), since $\mathscr{B}$ is a basis of $\mathbb{R}^k$. $\square$

Definition 5.6. Two bases $\mathscr{A}$ and $\mathscr{B}$ of $\mathbb{R}^k$ are called *compatible* if

$$\mathrm{int}(\mathrm{S}(\mathscr{A})) \cap \mathrm{int}(\mathrm{S}(\mathscr{B})) \neq \emptyset.$$

Otherwise $\mathscr{A}$ and $\mathscr{B}$ are called *incompatible*. (Cf. Figs. 5–7!)

Clearly, compatibility is no equivalence relation on the class of all bases of $\mathbb{R}^k$, since it is not transitive (cf. Figs. 5–7). But at least we get:

Corollary 5.7. *Let $\mathscr{A}$ and $\mathscr{B}$ be bases of $\mathbb{R}^k$. Then we have*:

(a) *$\mathscr{A}$ and $\mathscr{A}$ are compatible.*
(b) *$\mathscr{A}$ and $\mathscr{B}$ compatible $\Rightarrow$ $\mathscr{B}$ and $\mathscr{A}$ compatible.*

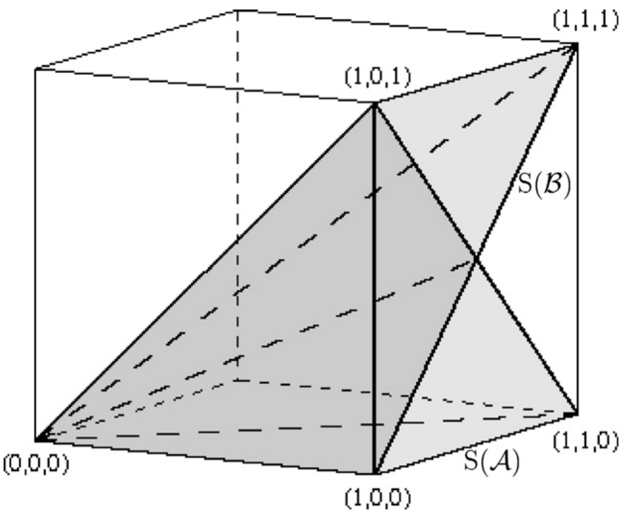

Fig. 5. An example of two compatible 0/1-bases for $k = 3$. Let $\mathscr{A}$ and $\mathscr{B}$ consist of the following 0/1-vectors:

$\mathscr{A}$	$\mathscr{B}$
$(1,0,0)$	$(1,0,0)$
$(1,1,0)$	$(1,0,1)$
$(1,0,1)$	$(1,1,1)$.

The bases $\mathscr{A}$ and $\mathscr{B}$ of $\mathbb{R}^3$ are compatible, since the interiors of the (3-dimensional) simplices $\mathrm{S}(\mathscr{A}) = \mathrm{conv}(\{0\} \cup \mathscr{A})$ and $\mathrm{S}(\mathscr{B}) = \mathrm{conv}(\{0\} \cup \mathscr{B})$ have common points.

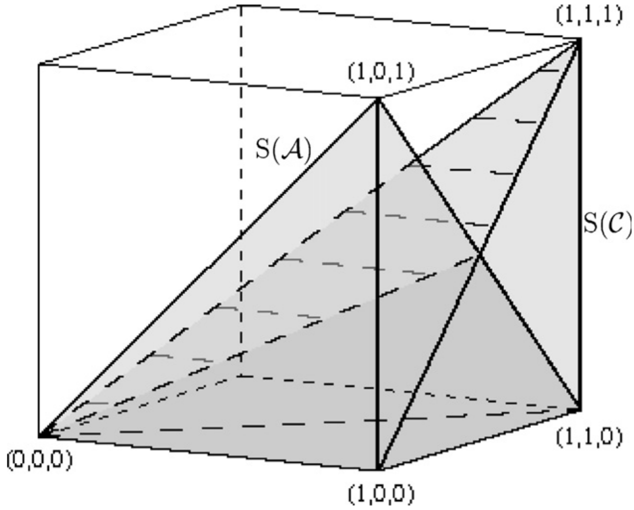

Fig. 6. An example of two compatible 0/1-bases for $k = 3$. Let $\mathscr{A}$ (cf. Fig. 5) and $\mathscr{C}$ consist of the following 0/1-vectors:

$\mathscr{A}$	$\mathscr{C}$
$(1,0,0)$	$(1,0,0)$
$(1,1,0)$	$(1,1,0)$
$(1,0,1)$	$(1,1,1)$.

The bases $\mathscr{A}$ and $\mathscr{C}$ of $\mathbb{R}^3$ are compatible, since the interiors of the (3-dimensional) simplices $\mathrm{S}(\mathscr{A})$ and $\mathrm{S}(\mathscr{C})$ have common points.

Proof. (a) is a consequence of Corollary 5.5(b), and (b) is trivial. □

Lemma 5.8. *Let $\mathscr{A}_{\geqslant} \subseteq \mathbb{R}^k$ be finite and $L\colon \mathscr{A}_{\geqslant} \to \mathbb{R}$. Then there exists an injective map*

$$\mathrm{choice}_L\colon \mathscr{E}(\mathscr{M}(L)) \to \mathrm{BAS}(\mathscr{A}_{\geqslant})$$

such that the image $\mathrm{choice}_L(\mathscr{E}(\mathscr{M}(L)))$ *is a set of pairwise incompatible bases of $\mathbb{R}^k$.*

Proof. Let $\mathscr{A}_{\geqslant} \subseteq \mathbb{R}^k$ be finite, and let $L\colon \mathscr{A}_{\geqslant} \to \mathbb{R}$. By Corollary 3.7(b), for every $x \in \mathscr{E}(\mathscr{M}(L))$ there exists a $\mathscr{B}(x) \in \mathrm{BAS}(\mathscr{A}_{\geqslant})$ such that $x \in \mathscr{M}(L, \mathscr{B}(x))$. Define

$$\mathrm{choice}_L : \mathscr{E}(\mathscr{M}(L)) \to \mathrm{BAS}(\mathscr{A}_{\geqslant}), \quad x \mapsto \mathscr{B}(x).$$

From Corollary 5.5(b) and Lemma 5.3 we receive

$$\forall x, y \in \mathscr{E}(\mathscr{M}(L)) \cdot (\mathscr{B}(x) \text{ and } \mathscr{B}(y) \text{ compatible} \Rightarrow x = y),$$

hence

$$\mathrm{choice}_L(\mathscr{E}(\mathscr{M}(L))) = \{\mathscr{B}(x) | x \in \mathscr{E}(\mathscr{M}(L))\}$$

is a set of pairwise incompatible bases of $\mathbb{R}^k$. By Corollary 5.7(a) choice_L is also injective. □

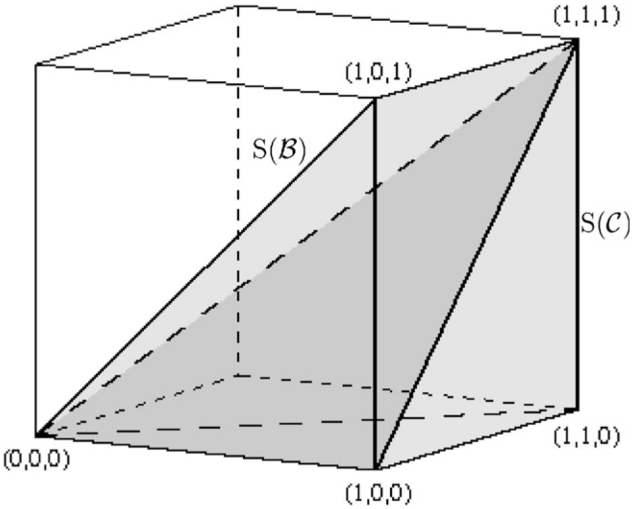

Fig. 7. An example of two incompatible 0/1-bases for $k = 3$. Let $\mathscr{B}$ (cf. Fig. 5) and $\mathscr{C}$ (cf. Fig. 6) consist of the following 0/1-vectors:

$\mathscr{B}$	$\mathscr{C}$
(1, 0, 0)	(1, 0, 0)
(1, 0, 1)	(1, 1, 0)
(1, 1, 1)	(1, 1, 1).

The bases $\mathscr{B}$ and $\mathscr{C}$ of $\mathbb{R}^3$ are incompatible, since the interiors (!) of the (3-dimensional) simplices $\mathrm{S}(\mathscr{B})$ and $\mathrm{S}(\mathscr{C})$ do not have common points.

On the one hand, in steps 1 and 2 we were not concerned with the specific properties of 0/1-bases. This alters in *step 3*, which has Lemma 5.12 as its goal.

On the other hand, in step 3 it will not be necessary to deal with the lower bounds L like above: We will now just consider 0/1-bases without regarding their roles defining vertices of polyhedra.

A simple fact is (cf. Figs. 5–7):

Proposition 5.9. *For every 0/1-basis* $\mathscr{B}$ *of* $\mathbb{R}^k$ *we have* $\mathrm{S}(\mathscr{B}) \subseteq [0;1]^k$.

Proof. The claim follows from $\{0\} \cup \mathscr{B} \subseteq \{0,1\}^k$, which implies $\mathrm{S}(\mathscr{B}) = \mathrm{conv}(\{0\} \cup \mathscr{B}) \subseteq \mathrm{conv}(\{0,1\}^k) = [0;1]^k$. $\square$

Definition 5.10. Let Vol_k be the *k-dimensional volume*, i.e., the Lebesgue measure on the Borel sets of $\mathbb{R}^k$.

Lemma 5.11. *For every 0/1-basis*[1] $\mathscr{B}$ of $\mathbb{R}^k$ *we have* $\mathrm{Vol}_k(\mathrm{S}(\mathscr{B})) \geqslant \frac{1}{k!}$.

[1] The statement is also true for bases, in which all the vectors have only integer components.

Proof. For every basis $\mathscr{B} = \{b_1, \ldots, b_k\}$ of $\mathbb{R}^k$ there is a well-known formula for the volume of the simplex $\mathrm{S}(\mathscr{B})$ (e.g., see [6], p. 374):

$$\mathrm{Vol}_k(\mathrm{S}(\mathscr{B})) = \frac{1}{k!} \cdot |\det \bar{\mathscr{B}}|,$$

where

$$\bar{\mathscr{B}} := (b_1 \ldots b_k),$$

if we consider $b_1, \ldots, b_k$ as column vectors. Now let $\mathscr{B}$ especially be a *0/1*-basis of $\mathbb{R}^k$. Since then $\bar{\mathscr{B}}$ is a matrix with integer entries, its determinant

$$d := \det \bar{\mathscr{B}}$$

is also an integer. But $d \neq 0$, since $\mathscr{B}$ is a basis of $\mathbb{R}^k$. Hence $|d| \geqslant 1$. □

Lemma 5.12. *Every set of pairwise incompatible 0/1-bases of* $\mathbb{R}^k$ *has at most* $k!$ *elements.*

Proof. Let $\mathscr{B}_1, \ldots, \mathscr{B}_n$ be pairwise incompatible 0/1-bases of $\mathbb{R}^k$. Then we have:

$$\frac{n}{k!} = \sum_{i=1}^{n} \frac{1}{k!} \overset{5.11}{\leqslant} \sum_{i=1}^{n} \mathrm{Vol}_k(\mathrm{S}(\mathscr{B}_i)) = \sum_{i=1}^{n} \mathrm{Vol}_k(\mathrm{int}(\mathrm{S}(\mathscr{B}_i))) \overset{(*)}{=} \mathrm{Vol}_k\left(\bigcup_{i=1}^{n} \mathrm{int}(\mathrm{S}(\mathscr{B}_i))\right)$$
$$\leqslant \mathrm{Vol}_k\left(\bigcup_{i=1}^{n} \mathrm{S}(\mathscr{B}_i)\right) \overset{5.9}{\leqslant} \mathrm{Vol}_k([0;1]^k) = 1,$$

where $(*)$ follows from the assumption that $\mathscr{B}_1, \ldots, \mathscr{B}_n$ are pairwise incompatible. □

Now, as our *fourth and last step* we just have to combine the results of step 2 and step 3:

Theorem 5.13. (Main theorem). $\psi(\mathscr{P}(\Omega_k)) \leqslant k!$.

Proof. Let $L\colon \mathscr{P}(\Omega_k) \to \mathbb{R}$. According to the definition of the psi-function (3.8(b)) we have to prove that

$$|\mathscr{E}(\mathscr{M}(L))| \leqslant k!.$$

By Lemma 5.8 there exists an injective map

$$\mathrm{choice}_L : \mathscr{E}(\mathscr{M}(L)) \to \mathrm{BAS}(\mathscr{P}(\Omega_k))$$

such that $\mathrm{choice}_L(\mathscr{E}(\mathscr{M}(L)))$ is a set of pairwise incompatible 0/1-bases of $\mathbb{R}^k$. Hence

$$|\mathscr{E}(\mathscr{M}(L))| = |\mathrm{choice}_L(\mathscr{E}(\mathscr{M}(L)))| \leqslant k!,$$

where the inequality is a consequence of Lemma 5.12. □

We conclude (cf. the definitions of (WAC2), (WEC), and M_k^F in Section 4):

Corollary 5.14.

(a) (WAC2) *is true.*
(b) (WEC) *is true.*
(c) $M_k^F = k!$ *for all* $k \geqslant 1$.

Proof.

(a) This is the statement of the Main theorem.
(b) Follows from (a) and Lemma 4.2.
(c) Follows from (b) and Lemma 4.1. □

6. Concluding remarks

How many extreme points structures can have, is an important question within the theory of interval probability. For example, it is often necessary to minimize or maximize linear functionals subject to structures, and therefore the complexity of corresponding algorithms could be estimated adequately.

We here computed the smallest upper bound for the number of extreme points for the class of coherent probabilities, and, indeed, for each kind of interval probability in finite spaces.

It is a remarkable fact that this bound ($k!$ for sample spaces with k elements) does not exceed the well-known bound for the – very small – subclass of 2-monotone capacities.

The technique for proving this result here was abstract: by considering properties of the dual space of $\mathbb{R}^k$. A constructive proof instead would – more or less – presuppose a pattern for describing the huge variety of forms structure vertices of coherent probabilities can possess (even in the case of uniform probabilities according to (10); see the pictures in [8], Chapter 5).

The flexibility of coherent probabilities in comparison with 2-monotone capacities induces that the pattern (8) (or (9)), generated by permutations of $\{1,\ldots,k\}$ and only valid for 2-monotone capacities, is far away from an appropriate classification.

Nevertheless, having such a classification would be extremely interesting for characterizing coherent probability. In particular, it would lead to a description of its (existing!) natural division into convex subclasses; the class of 2-monotone capacities is only one of them.

Acknowledgements

I thank Thomas Augustin, Peter Gritzmann, Kurt Weichselberger, and anonymous referees for their valuable comments. Especially I thank Franz Ferschl – as well for structural pieces of advice as for algebraic and geometric stimuli.

References

[1] A. Chateauneuf, J.-Y. Jaffray, Some characterizations of lower probabilities and other monotone capacities through the use of Möbius inversion, Mathematical Social Sciences 17 (1989) 263–283.
[2] P.J. Huber, V. Strassen, Minimax tests and the Neyman–Pearson lemma for capacities, Annals of Statistics 1 (1973) 251–263, Correction: 2 (1974) 223–224.
[3] I. Levi, The Enterprise of Knowledge, MIT Press, Cambridge, MA, 1980.
[4] M.W. Padberg, Linear Optimization and Extensions, Springer, Berlin, Heidelberg, New York, 1995.
[5] S. Shapley, Cores of convex games, International Journal of Game Theory 1 (1971) 12–26.
[6] K.T. Smith, Primer of Modern Analysis, Springer, New York, Berlin, Heidelberg, 1983.
[7] P. Walley, Statistical Reasoning with Imprecise Probabilities, Chapman and Hall, London, New York, 1991.

[8] A. Wallner, Beiträge zur Theorie der Intervallwahrscheinlichkeit – Der Blick über Kolmogorov und Choquet hinaus, Dr. Kovač, Hamburg, 2002.
[9] K. Weichselberger, Elementare Grundbegriffe einer allgemeineren Wahrscheinlichkeitsrechnung I – Intervallwahrscheinlichkeit als umfassendes Konzept, in cooperation with T. Augustin, A. Wallner, Physica, Heidelberg, 2001.